

基于 CNN 的珠宝分类系统设计与实现

周涵庆 柯圆圆*

(江汉大学人工智能学院, 湖北 武汉 430056)

摘要 随着神经网络的发展, 生活中越来越多的地方涉及到神经网络, 例如, 通过神经网络进行文字识别、图像检索、人脸识别等。本文基于 CNN 设计了一个珠宝识别系统, 以 CNN 为核心, 使用七类珠宝图片作为数据集, 以 Pyqt5 构建图形化界面, 经过训练可以准确识别珠宝种类。

关键词 珠宝识别 神经网络 CNN

中图分类号: P619.28

文献标识码: A

文章编号: 1007-0745(2022)09-0121-03

随着生活水平的提高和人民对精神层面的需求, 购买珠宝的消费者也越来越多。大多数消费者选购珠宝时会认为品牌店的珠宝有质量保证, 因此会去珠宝品牌店购买。但是一段时间以来, 珠宝市场上打折风愈刮愈烈, 从而引发了一场消费者对珠宝经营者的信任危机。因此, 我们迫切地需要对珠宝饰品价格进行透明的可信的价格评估。传统的珠宝评估完全由人来完成, 这使得用户并不知道专家是如何评估的, 形成了信息不对称的局面。珠宝价格评估通过计算机来完成时就可以消除这种信息不对称。其逻辑如下: (1) 利用用户输入的珠宝图片。(2) 计算机将图片与后台大型数据库中的图片进行检索。(3) 输出相似度在百分之八十以上的珠宝图片与相关信息。本文基于第二点进行改进, 在检索之前将数据库中珠宝图片进行分类存储, 那么图片检索时就可以只检索数据库中相应类别的图片, 而不必检索整个数据库。基于此, 本文提出基于 CNN 的珠宝分类系统进行图片识别。

1 CNN 相关操作详解

基于深度学习的方法相比于传统的方法提取特征点更快、更准确, 因此越来越多的学者将深度学习引入图像识别领域^[1]。卷积神经网络(CNN), 一种用于处理图像的神经网络, 从图像中获取输入^[2], 并从图像中提取特征, 通过 CNN 提取的特征适合于珠宝图片的分类^[3], 并提供可学习的参数以有效地进行分类, 从而使图像分类任务的准确率和准确率有显著提升^[4]。有关 CNN 更多的研究内容见参考文献^[5-8]。

1.1 卷积

卷积神经网络中最基础的操作也是最重要的操作, 也就是卷积操作, 使用卷积核(过滤器)从图像中提取

特征, 不同的过滤器可以从图像中提取不同的特征。提取特征的过程就是卷积运算, 卷积核与图像像素矩阵局部做点积运算。卷积核一般大小采用 3×3 。卷积运算中一些参数有: 步幅(stride), 步幅也就是与卷积核进行运算的像素矩阵上的矩阵窗口移动的单位。步幅分为水平步幅与垂直步幅, 一般默认垂直步幅为 1, 水平步幅可以为 1, 也可以为 2。步幅一般是指水平步幅。此外, 还考虑到有时候进行卷积时不希望图像大小发生变化, 或者当图片的重要信息在图片边缘时, 为了保护边缘信息不丢失, 此时采用零填充(zero-padding)的方式来保护边缘信息不丢失或者维持图像大小不变, 所谓零填充即在图像矩阵的四周补上一圈或者几圈的 0。

例如, 当用一个 3×3 的卷积核对一个 6×6 的像素矩阵进行卷积, 第一次运算时, 第一行到第三行, 第一列到第三列组成的 3×3 矩阵窗口与卷积核进行点乘, 输出结果为 17, 随后矩阵窗口以水平步幅为 1 向右滑动 1 个位置, 将新的矩阵窗口与卷积核再次进行点乘运算, 输出结果为 12。当完成水平方向的滑动到最后三列时, 垂直位移一个单位, 进行第二行到第四行, 第一列到第三列的卷积, 并不断进行水平方向的位移。卷积核不断滑过像素矩阵, 最终得到特征图。

一个 $n \times n$ 的像素矩阵与一个 $m \times m$ 的卷积核做卷积运算, 通过公式 $((n-m)/stride)+1$ 可以计算出特征图的维度, 但是当输入图像的像素矩阵为 32×32 , 卷积核为 4×4 , stride 为 3 时, $(32-4)/3+1$ 不是整数, 这时候可以引入零填充(padding)的方式, 即在图像像素矩阵添一圈 0, 这样输出特征图的维度就变成了 $(32+1*2-4)/3+1=11$, 在神经网络中必须要掌握每一层输入输出的维度。

*本文通讯作者, E-mail: keyy086@126.com.

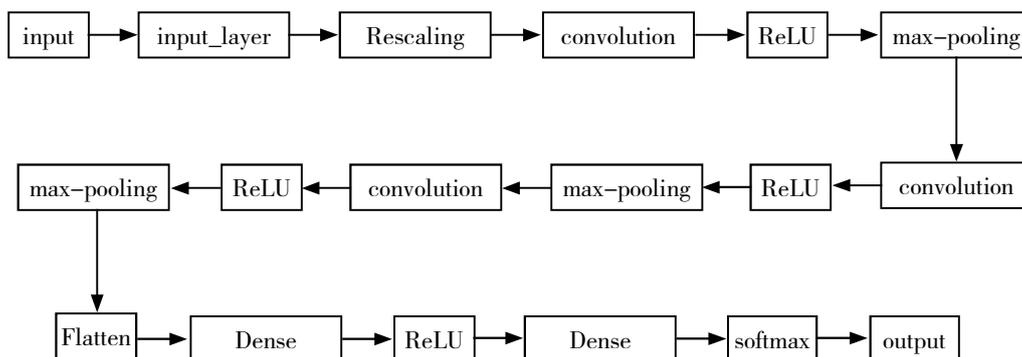


图1 本文 CNN 结构图

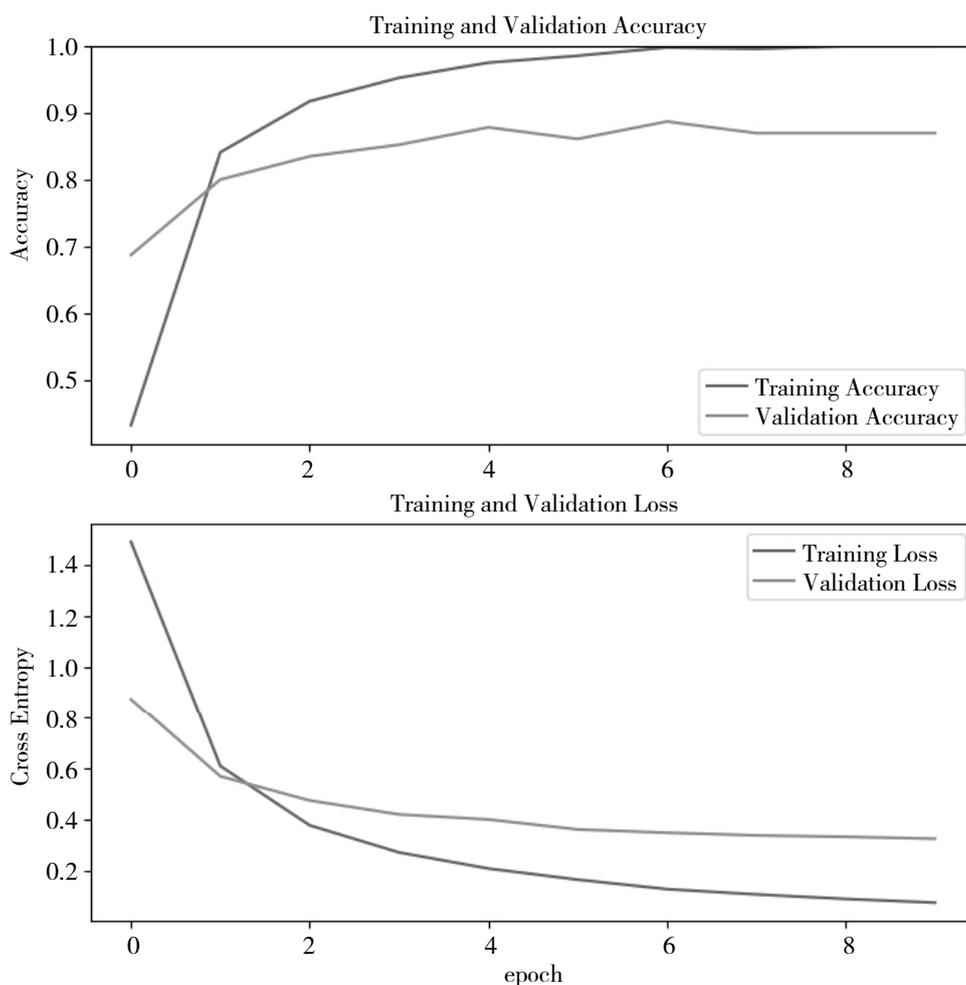


图2 模型训练过程图的准确率和损失率

一般而言，用户输入的都是彩色照片，彩色照片的深度为3，也就是RGB三通道，每一个像素点由3个值组成，我们的卷积核的通道应该和输入图片的通道一致，例如照片的图像像素矩阵为 $32 \times 32 \times 3$ ，卷积核可以设置为 $3 \times 3 \times 3$ ，将卷积核的每一层和像素矩阵

的每一层做卷积运算，最后再组成一个深度为3的输出，这里 stride 设置为1，padding 为0，所以输出的 shape 为 $30 \times 30 \times 3$ 。

卷积的时候是用多个卷积核完成的，一般经过卷积之后的特征图的通道为卷积核的数量。

1.2 激活与池化

卷积完成之后,下一步引入非线性激活函数,给神经网络加入一些非线性因素,使得神经网络可以更好地解决较为复杂的问题,激活函数有很多种,本文使用ReLU,对于输入的负值,输出全为0,对于正值,原样输出,优点是收敛快,求梯度简单。

激活的下一步就是池化,池化可以将实现精简特征图缩小^[9],降低计算量同时又保留其中的重要信息,池化分为平均池化和最大池化,本文选择最大池化(max-pooling,也叫下采样)保留了每一个小块内的最大值,相当于保留了这一块最佳的结果。对于一个 $n \times n$ 的像素矩阵,保存每个 2×2 的矩阵内最大值来替代这个 2×2 矩阵,就可以使得其规模缩小至 $(n/2) \times (n/2)$ 。

2 CNN 结构

上面我们介绍了卷积、激活与池化,而一个神经网络就是不断地进行卷积、激活与池化的过程。图1所示为本文CNN结构图。

其中,Rescaling为对模型做归一化处理,将像素矩阵值在0-255处理到0-1之间。Convolution为卷积操作,激活函数使用ReLU,随后进行max-pooling,再进行两次卷积、激活、池化。Flatten将二维输出展开成一维,随后接入全连接层。损失函数使用交叉熵损失函数。

数据集通过爬虫免费获取,共七类珠宝图片,分别为:“蓝宝石戒指”“蓝宝石项链”“钻石戒指”“银吊坠”“黄金吊坠”“黄金戒指”“黄金手镯”,共599张图片。按照0.8:0.2的比例划分为训练集与测试集,在划分数据集之前切记对原始数据集进行检查,因为爬虫获取的图片可能存在坏图,这会导致网络训练过程异常。

3 网络实现

训练轮数epoch设为10轮,训练完成后得到模型训练过程图如图2。

从模型训练过程图2可以看到随着训练轮数的增加训练集和测试集上的准确率均在增加,随着训练轮数的增加训练集和测试集上的损失函数均在减小。

实验结果为蓝宝石戒指的预测成功率为百分之九十四,蓝宝石项链的预测成功率为百分之八十二,钻石戒指的预测成功率为百分之八十八,银吊坠的预测成功率为百分之八十七,黄金戒指的预测成功率为百分之八十三,对黄金吊坠的分类达到百分百的成功率。其中对黄金手镯的预测成功率仅为百分之七十二,

并且对于黄金手镯的预测将有百分之二十八的概率错误的识别为黄金戒指。对于黄金戒指的预测有百分之十七的概率预测为黄金手镯。因此在进行数据库的珠宝图片分类时,需要对分类结果进行核查。

实验结果表明最终可以实现对珠宝图片的自动化识别。以提前对图像数据库进行分类,在图像检索的过程中,识别出具体珠宝类别,就可以指定检索某一类图像,而不用对全部图像进行检索,大大节省了计算机资源。

参考文献:

- [1] 张林鹏,汪西原,李强.基于双池化特征加权结构CNN的图像分类[J].计算机与现代化,2021(11):67-71,88.
- [2] 刘斌,龙健宁,程方毅,等.基于卷积神经网络的物流货物图像分类研究[J].机电工程技术,2021,50(12):79-82,175.
- [3] 林丽惠,魏毅,潘俊虹.基于卷积神经网络的武夷岩茶叶片分类方法[J].宁德师范学院学报(自然科学版),2021,33(04):363-369.
- [4] 李彦甫,范习健,杨绪兵,等.基于自注意力卷积神经网络的遥感图像分类[J].北京林业大学学报,2021,43(10):81-88.
- [5] Shan Wang, Yue Wang, Qilong Zhao, et al. Exploring Classification Capability of CNN Features[C]//Abstracts of the 7th International Conference of Pioneering Computer Scientists, Engineers and Educators (ICPCSEE 2021) Part I. Abstracts of the 7th International Conference of Pioneering Computer Scientists, 2021:86-88.
- [6] Zhu Anguo, Liu Longjun, Hou Wenxuan, et al. HSC: Leveraging horizontal shortcut connections for improving accuracy and computational efficiency of lightweight CNN[M]. Elsevier BV, 2021:141-145.
- [7] Tangbo Bai, Jianwei Yang, Guiyang Xu, et al. An optimized railway fastener detection method based on modified Faster R-CNN[M]. Elsevier BV, 2021:109742.
- [8] Euphrem Mugisha Rwagasore, Xiong Zhang, Kaifang Gao, et al. Convolutional neural network for sapphire ingots defect detection and classification[M]. Optical Materials, 2021:111292.
- [9] 田晟,宋霖.基于CNN和Bagging集成的交通标志识别[J/OL].广西师范大学学报(自然科学版),2022-01-16:1-13.